

# X-ray Cluster Constraints on Non-Gaussianity

Sarah Shandera,<sup>a,f</sup> Adam Mantz,<sup>b,f</sup> David Rapetti,<sup>c</sup> Steven W. Allen<sup>d,e</sup>

<sup>a</sup>Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

<sup>b</sup>Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA

<sup>c</sup>Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark

<sup>d</sup>Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA

<sup>e</sup>SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

<sup>f</sup>Corresponding authors

E-mail: [shandera@gravity.psu.edu](mailto:shandera@gravity.psu.edu), [amantz@kicp.uchicago.edu](mailto:amantz@kicp.uchicago.edu),  
[drapetti@dark-cosmology.dk](mailto:drapetti@dark-cosmology.dk), [swa@stanford.edu](mailto:swa@stanford.edu)

**Abstract.** We report constraints on primordial non-Gaussianity from the abundance of X-ray detected clusters. Our analytic prescription for adding non-Gaussianity to the cluster mass function takes into account moments beyond the skewness, and we demonstrate that those moments should not be ignored in most analyses of cluster data. We constrain the amplitude of the skewness for two scenarios that have different overall levels of non-Gaussianity, characterized by how amplitudes of higher cumulants scale with the skewness. We find that current data can constrain these one-parameter non-Gaussian models at a useful level, but are not sensitive to adding further details of the corresponding inflation scenarios. Combining cluster data with Cosmic Microwave Background constraints on the cosmology and power spectrum amplitude, we find the dimensionless skewness to be  $10^3 \mathcal{M}_3 = -1^{+24}_{-28}$  for one of our scaling scenarios, and  $10^3 \mathcal{M}_3 = -4 \pm 7$  for the other. These are the first constraints on non-Gaussianity from Large Scale Structure that can be usefully applied to any model of primordial non-Gaussianity. The former constraint, when applied to the standard local ansatz (where the  $n$ -th cumulant scales as  $\mathcal{M}_n \propto \mathcal{M}_3^{n-2}$ ), corresponds to  $f_{\text{NL}}^{\text{local}} = -3^{+78}_{-91}$ . When applied to a model with a local-shape bispectrum but higher cumulants that scale as  $\mathcal{M}_n \propto \mathcal{M}_3^{n/3}$  (the second scaling scenario), the amplitude of the local-shape bispectrum is constrained to be  $f_{\text{NL}}^{\text{local}*} = -14^{+22}_{-21}$ . For this second scaling (which occurs in various well-motivated models of inflation), we also obtain strong constraints on the equilateral and orthogonal shapes of the bispectrum,  $f_{\text{NL}}^{\text{equil}} = -52^{+85}_{-79}$  and  $f_{\text{NL}}^{\text{orth}} = 63^{+97}_{-104}$ . This sensitivity implies that cluster counts could be used to distinguish qualitatively different models for the primordial fluctuations that have identical bispectra.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The effect of Primordial non-Gaussianity on object number counts</b>	<b>4</b>
2.1	Parametrizing the level of non-Gaussianity	4
2.2	The mass function in terms of $\mathcal{M}_3$ and the scaling of higher moments	6
<b>3</b>	<b>Two parameter extensions</b>	<b>7</b>
<b>4</b>	<b>Application to galaxy cluster data</b>	<b>8</b>
<b>5</b>	<b>Comparison with the literature</b>	<b>13</b>
<b>6</b>	<b>Conclusions</b>	<b>17</b>
<b>A</b>	<b>Non-Gaussian mass functions</b>	<b>18</b>
A.1	Other non-Gaussian mass functions	20

---

## 1 Introduction

The Large Scale Structure of the late Universe depends on a rich array of physics: the spectrum of primordial curvature inhomogeneities, the cosmological evolution of the Universe, the rules governing the growth of structure, and particle physics, chemistry and thermodynamics within individual stars, galaxies, and galaxy clusters. Extracting details of the primordial fluctuations is necessarily a difficult problem, but fortunately there are several complementary observables available to us. In this work, we use measurements of the mass and redshift distribution (the mass function) of a sample of galaxy clusters to constrain primordial non-Gaussianity. We demonstrate that this is a complementary probe to the Cosmic Microwave Background (CMB) bispectrum and the halo bias because it is sensitive to different aspects of the primordial non-Gaussianity.

Most work on cosmological constraints from clusters has focused on dark energy [1–3]. Here we apply the substantial progress made on characterizing the mass–observable relations in that context to study primordial non-Gaussianity. We use 237 X-ray bright clusters detected in the ROSAT All-Sky Survey [4] and the analysis techniques of Mantz et al. [5] to investigate two one-parameter models for the non-Gaussian curvature perturbations and four two-parameter models. The clusters in this sample have redshifts up to  $z = 0.5$  and masses of order  $10^{14}$ – $10^{15} M_\odot$ . With the semi-analytic, non-Gaussian mass function extended to include terms beyond the skewness, we find constraints that are completely consistent with Gaussian statistics for the primordial fluctuations. However, contrary to some of the expectations in the literature [6–8], we find error bars small enough to indicate that cluster counts can provide complementary information to current constraints from the CMB bispectrum [9, 10] and from the galaxy bias [11–14]. Our results are consistent with but tighter than those recently obtained using clusters detected by the South Pole Telescope (SPT) [15, 16] and clusters selected from the Sloan Digital Sky Survey (SDSS) [17].

Statistics of the primordial fluctuations beyond the homogeneous and isotropic power spectrum are an extremely important source of information about the very early Universe

– so much so that pursuing limits on non-Gaussianity down to the minimal levels expected from single field slow-roll inflation is an important task. Galaxy clusters form from rare primordial over-densities, with more massive and higher redshift clusters tracing rarer initial fluctuations, and so their abundance is sensitive to non-Gaussianity in the primordial inhomogeneities. Constraints from cluster number counts are complementary to other probes of non-Gaussianity in three ways: they probe smaller scales than the CMB or galaxy bias do (cluster constraints are at  $k \approx 0.1\text{--}0.5 h/\text{Mpc}$ ), they are sensitive to *any* non-Gaussianity (including any shape for the primordial bispectrum), and they are sensitive to higher order cumulants of the probability distribution of the primordial inhomogeneities.

The non-Gaussian statistics of the primordial perturbations are completely described by the set of correlation functions  $\langle \Phi(\vec{k}_1)\Phi(\vec{k}_2)\dots\Phi(\vec{k}_n) \rangle_c$ , where  $\Phi(k)$  is the Bardeen potential in momentum space, and  $n$  runs from 3 to infinity. The subscript  $c$  stands for ‘connected’ and picks out the parts of the correlations that vanish for a Gaussian field. Clearly, a single parameter cannot describe this series of functions. However, if the field is weakly non-Gaussian, the three-point function is likely to generate the strongest observational signal, so many non-Gaussian models are classified by naming the configuration of the three-point function and its amplitude. The Wilkinson Microwave Anisotropy Probe (WMAP) satellite, for example, reports constraints on the parameters  $f_{\text{NL}}^{\text{local}}$ ,  $f_{\text{NL}}^{\text{equil}}$ ,  $f_{\text{NL}}^{\text{orth}}$ , where the labels ‘local’, ‘equilateral’, and ‘orthogonal’ refer to specific, scale-independent functional forms assumed for the three-point correlation. Translation invariance requires that the three-point correlation has a factor  $\delta_D^3(\vec{k}_1 + \vec{k}_2 + \vec{k}_3)$  (a Dirac delta function), similar to the term  $\delta_D^3(\vec{k}_1 + \vec{k}_2)$  in the power spectrum. The three-point correlation is then just a function of two independent momenta and is called the bispectrum.

The local, equilateral and orthogonal bispectra are shown in Eq.(2.3) below. Interestingly, though, object number counts are not sensitive to the details of the bispectrum’s momentum dependence. Instead, only the integrated moments of the smoothed density fluctuations  $\delta_R$  are relevant. For example,

$$\langle \delta_R^3 \rangle = \int \frac{d^3\vec{k}_1}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} \int \frac{d^3\vec{k}_3}{(2\pi)^3} M(k_1, R, z) M(k_2, R, z) M(k_3, R, z) \langle \Phi(\vec{k}_1)\Phi(\vec{k}_2)\Phi(\vec{k}_3) \rangle_c \quad (1.1)$$

where the terms  $M(k_i, R, z)$  contain a window function, the corresponding factors from the Poisson equation, the transfer function and the growth factor converting the linear perturbation in the gravitational potential to the smoothed density perturbation. (The full expressions for these quantities can be found in Appendix A.) We characterize the non-Gaussianity by the dimensionless ratios of the cumulants of the density field

$$\mathcal{M}_{n,R} = \frac{\langle \delta_R^n \rangle_c}{\langle \delta_R^2 \rangle^{n/2}} \quad (1.2)$$

which are by construction redshift independent and nearly independent of the smoothing scale,  $R$ , if the primordial bispectrum is scale independent.<sup>1</sup>

When cumulants beyond the skewness (correlations beyond the bispectrum) are relevant, a one-parameter model is only useful if we can use it to specify the amplitude of *all* the correlations. In this paper we use  $\mathcal{M}_3$  and a choice for how higher moments scale with  $\mathcal{M}_3$  to describe non-Gaussian fluctuations. The scalings we consider are motivated by particle

---

<sup>1</sup>Scale independence means that the bispectrum (e.g., those in Eq.(2.3)) contains no length scale other than the factors  $k_i^{-1}$  in the  $P(k_i)$  terms.

physics models of inflation, and our constraints on the total dimensionless skewness can always be re-written in terms of a particular bispectrum using Eq.(1.1) and Eq.(1.2).

Most previous work on the utility of cluster counts to constrain non-Gaussianity has focused on the local ansatz [18, 19], where one assumes that the non-Gaussian field  $\Phi(x)$  is a simple, local transformation of a Gaussian field  $\Phi_G(x)$ :

$$\Phi(x) = \Phi_G(x) + f_{\text{NL}}^{\text{local}}[\Phi_G(x)^2 - \langle \Phi_G(x)^2 \rangle]. \quad (1.3)$$

In this useful model,  $f_{\text{NL}}^{\text{local}}$  is the single parameter that all correlation functions depend on, and the cumulants scale as  $(f_{\text{NL}}^{\text{local}})^{n-2}$ . Non-Gaussianity of the local type has a bispectrum that most strongly correlates Fourier modes of very different wavelengths. This particular mode coupling generates strong signals in other large scale structure observables – most notably introducing a scale dependence in the bias of dark matter halos, luminous red galaxies and quasars [11, 14, 20]. For this reason, papers that have analyzed the potential for future surveys to constrain non-Gaussianity have largely focused on non-Gaussianity captured by the local ansatz only, and on the superior constraints from the bias compared to number counts. The bias may allow us to probe  $\Delta f_{\text{NL}}^{\text{local}} \sim \mathcal{O}(1)$  in the near future [21, 22], although this optimism is still subject to a full understanding of the relevant systematics [14, 23, 24]. Regardless, the motivation for looking at cluster number counts to constrain a scale-invariant, local ansatz is certainly weak. However, single field models of inflation cannot produce large local non-Gaussianity over a wide range of scales [25]. Since non-Gaussianity that does not strongly correlate modes of very different wavelength is *not* particularly detectable in the galaxy bias [13, 26, 27], that measure alone tests only a subset of viable inflation models. Furthermore, number count constraints are sensitive to higher order correlation functions. As we will demonstrate here, number counts can distinguish between scenarios with indistinguishable bispectra that nonetheless are generated by qualitatively different inflationary physics. That is, a model may have a local-shape bispectrum but higher order moments that do not scale as those from the standard local ansatz do (e.g., [28]). In this case, number counts or other measurements sensitive to higher moments will provide complementary information to the halo bias constraints. For the rest of the paper, we will assume that  $f_{\text{NL}}^{\text{local}}$  is a parameter measured from the bispectrum alone (see Eq.(2.3) below), which does not imply the entire series of correlations that Eq.(1.3) generates, unless we specify the local ansatz as our model. This use of  $f_{\text{NL}}^{\text{local}}$  is more in line with how it is observationally defined (e.g., in analyses of the CMB and halo bias).

This paper provides the first constraint on primordial non-Gaussianity from X-ray detected clusters, and the first large scale structure constraint that can be usefully applied to any model for primordial non-Gaussianity. In the next Section, we define our parametrization of the effects of primordial non-Gaussianity on cluster abundance, using a semi-analytic non-Gaussian mass function in terms of a single new parameter. In Section 3 we discuss several theoretically motivated extensions to two parameters. We present our results for the one and two parameter models in Section 4. In Section 5, we discuss how our results compare to previous results and forecasts in the literature, including the SPT constraints reported from two small samples of Sunyaev-Zel’dovich (SZ) detected clusters [15, 16] and those from a large sample of optically selected clusters [17], the SDSS maxBCG cluster catalogue [29]. We summarize in Section 6. Appendix A contains the details of our semi-analytic mass function prescriptions as well as others that exist in the literature.<sup>2</sup> Quoted error intervals are always 68.3% confidence, unless otherwise specified.

<sup>2</sup>We have shared some computer code helpful for evaluating our mass function; see Appendix A.

## 2 The effect of Primordial non-Gaussianity on object number counts

Our basic tool is a series expansion for the ratio of the non-Gaussian mass function to the Gaussian one. The expansion we use is based on a Press-Schechter model for halo formation applied to non-Gaussian probability distributions for the primordial fluctuations. A detailed derivation of the non-Gaussian mass function we use is given in Appendix A and was developed in [30–32]. The weakly non-Gaussian probability distributions that the mass function is based on are asymptotic expansions that deviate substantially from the actual probability density function (PDF) for sufficiently rare fluctuations. Fortunately, our cosmology is already sufficiently constrained to determine that the clusters in our sample do not lie in that regime. However, the clusters *are* sufficiently rare that truncating the expansion below at a single term (the skewness) is not sufficient to test the full range of models that are only as skewed as current CMB constraints allow.

We add non-Gaussianity to the cosmology by considering a mass function of the form

$$\left(\frac{dn}{dM}\right)_{\text{NG}} = \left(\frac{dn}{dM}\right)_{\text{T}, M_{300}} \left(\frac{n_{\text{NG}}}{n_{\text{G}}}\right)_{\text{Edgeworth}} \quad (2.1)$$

where the first term on the right hand side is the Gaussian mass function of Tinker et al. [33] for clusters identified as spheres containing a mean density 300 times that of the mean matter density of the Universe,  $300 \bar{\rho}_{\text{m}}(z)$ . The ratio of the non-Gaussian mass function to the Gaussian one will be given as a series expansion, defined below. This factor will be a function of mass, redshift, and parameters that characterize the amplitude of the non-Gaussianity, which we define next.

### 2.1 Parametrizing the level of non-Gaussianity

Since object number counts are not sensitive to the details of the momentum space correlations, we consider the dimensionless, connected moments (the cumulants, divided by the appropriate power of the amplitude of fluctuations) of the density fluctuations smoothed on a given scale  $R$ , as defined in Eq.(1.2). Most constraints on non-Gaussianity have so far been reported for a parameter that measures the size of the three-point correlation in momentum space, or bispectrum. This is an extremely useful first statistic because this correlation should be exactly zero if the fluctuations were exactly Gaussian. However, because the bispectrum is a function of two momenta, the non-Gaussian parameters most often quoted assume a shape for the bispectrum.

A generic homogeneous and isotropic bispectrum for the potential  $\Phi$  can be written as

$$\left\langle \Phi(\vec{k}_1) \Phi(\vec{k}_2) \Phi(\vec{k}_3) \right\rangle_c = (2\pi)^3 \delta_D^3(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B(k_1, k_2, k_3) \quad (2.2)$$

where the function  $B(k_1, k_2, k_3)$  determines the shape. Bispectra are colloquially named by the (triangle) configuration of the three momentum vectors that are most strongly correlated. To interpret our constraints on  $\mathcal{M}_3$  in terms of familiar bispectra, we consider the templates for ‘local’, ‘equilateral’ and ‘orthogonal’ bispectra:

$$\begin{aligned} B_{\text{local}} &= 2f_{\text{NL}}^{\text{local}}(P(k_1)P(k_2) + P(k_1)P(k_3) + P(k_2)P(k_3)) \\ B_{\text{equil}} &= 6f_{\text{NL}}^{\text{equil}}[-P(k_1)P(k_2) + 2 \text{ perm.} - 2(P(k_1)P(k_2)P(k_3))^{2/3} \\ &\quad + P(k_1)^{1/3}P(k_2)^{2/3}P(k_3) + 5 \text{ perm.}] \\ B_{\text{orth}} &= 6f_{\text{NL}}^{\text{orth}}[-3P(k_1)P(k_2) + 2 \text{ perm.} - 8(P(k_1)P(k_2)P(k_3))^{2/3} \\ &\quad + 3P(k_1)^{1/3}P(k_2)^{2/3}P(k_3) + 5 \text{ perm.}] \end{aligned} \quad (2.3)$$

where the power spectrum,  $P(k)$ , is defined from the two-point correlation function by

$$\langle \Phi(\vec{k}_1) \Phi(\vec{k}_2) \rangle = (2\pi)^3 \delta_D^3(\vec{k}_1 + \vec{k}_2) P(k_1) \equiv (2\pi)^3 \delta_D^3(\vec{k}_1 + \vec{k}_2) 2\pi^2 \frac{\Delta_\Phi^2(k_0)}{k^3} \left( \frac{k_1}{k_0} \right)^{n_s-1}. \quad (2.4)$$

In the best fit cosmology from the seven-year WMAP data, baryon acoustic oscillations and Hubble parameter measurements, the spectral index is  $n_s = 0.967$  [34, 35], and the amplitude is such that  $\sigma_8 = 0.81$ . Observationally, the parameter  $f_{\text{NL}}^{\text{local}}$  is typically measured by looking for a bispectrum of the form given in the first line of Eq.(2.3), which has weaker implications than the definition of *all* the correlation functions as in Eq.(1.3). The best current limits on the amplitudes of the bispectra in Eq.(2.3) come from the *Planck* Satellite maps of the CMB [10], which limit  $f_{\text{NL}}^{\text{local}} = 2.7 \pm 5.8$ ,  $f_{\text{NL}}^{\text{equil}} = -42 \pm 75$ , and  $f_{\text{NL}}^{\text{orth}} = -25 \pm 39$  at the 68.3% confidence level. Table 1 shows the value of  $\mathcal{M}_3$ , smoothed on a scale corresponding to  $10^{14} h^{-1} M_\odot$  halos ( $h$  is related to the Hubble parameter today,  $H_0 = 100h$  km/s/Mpc), for the local, equilateral, and orthogonal templates using the WMAP7 best fit cosmology.

**Table 1.** The conversions between the parameter  $\mathcal{M}_3$  and the amplitudes of particular bispectra  $f_{\text{NL}}$ . These numbers assume the WMAP7 best fit cosmology and change by at most a few percent if the best fit cosmologies from our analysis are used instead.

Shape	$\mathcal{M}_3$
Local	$0.00031 f_{\text{NL}}^{\text{local}}$
Equilateral	$0.000086 f_{\text{NL}}^{\text{equil}}$
Orthogonal	$-0.000062 f_{\text{NL}}^{\text{orth}}$

For some non-Gaussian scenarios (notably the local ansatz and typical single field models) the parameter  $\mathcal{M}_3$  is interchangeable with the  $f_{\text{NL}}^{\text{local}}$ ,  $f_{\text{NL}}^{\text{equil}}$ , or  $f_{\text{NL}}^{\text{orth}}$  as a description of the amplitude of the three point function *and* as a measure of the total non-Gaussianity for the entire series of higher order correlations. This is possible when the cumulants scale parametrically as

$$\mathcal{M}_n^{\text{h}} \propto \left( \mathcal{I} \mathcal{P}^{\frac{1}{2}} \right)^{n-2}, \quad (2.5)$$

where  $\mathcal{I}$  is proportional to the appropriate  $f_{\text{NL}}$  parameter. Although it is not needed for the simplest models, we use this more general notation, since it is useful for the two-parameter scenarios we introduce below. The superscript ‘h’ labels the scaling in Eq.(2.5), which we call *hierarchical*.

Mathematically, the higher order correlations could be nearly arbitrary, and it is only because we hope they have a common origin in some perturbation theory that it seems likely they are related. In this paper, we will contrast a second possible scaling, occurring in models where the scalar inflaton couples to a gauge field, that is much more non-Gaussian than the hierarchical scenario for a fixed value of the skewness [28, 31]. This scaling, which we call *feeder* (since it originates in models where fluctuations of a second field provide an extra source for the inflaton fluctuations<sup>3</sup>) is

$$\mathcal{M}_n^{\text{f}} \propto \mathcal{I}^n, \quad n \geq 3. \quad (2.6)$$

<sup>3</sup>The equation of motion for the inflaton fluctuations  $\delta\varphi$  is  $[\partial_t^2 - a^{-2}\nabla^2 + 3H\partial_t + m^2]\delta\varphi = J$ , where  $J$  is a source that depends on the quantum fluctuations of fields coupled to the inflaton [31].



Object number counts are sensitive to the value of the total skewness and to the scaling of higher moments, rather than any details of the momentum space correlations.

In addition to the dependence on a parameter like  $f_{\text{NL}}$ , the cumulants also have numerical coefficients that typically have to do with combinatorics. For example, beginning with Eq.(1.3), the bispectrum contains three terms linear in  $f_{\text{NL}}^{\text{local}}$ , each with two equivalent ways to take the expectation value of pairs of fields  $\Phi_{\text{G}}$ . We will choose the constants of proportionality equal to combinatoric factors for the moments that are generated in the local ansatz and a simple two-field extension that gives feeder scaling:<sup>4</sup>

$$\text{Hierarchical} \quad \mathcal{M}_n^{\text{h}} = n! 2^{n-3} \left( \frac{\mathcal{M}_3^{\text{h}}}{6} \right)^{n-2} \quad (2.7)$$

$$\text{Feeder} \quad \mathcal{M}_n^{\text{f}} = (n-1)! 2^{n-1} \left( \frac{\mathcal{M}_3^{\text{f}}}{8} \right)^{n/3}. \quad (2.8)$$

For a given scaling of the moments, we can determine a series expansion for the probability distribution and for the mass function that can be consistently truncated at some order in the moments.

For the single parameter scenarios, we report constraints in terms of the scaling assumed and the parameter  $\mathcal{M}_3$ , which can be compared with other constraints on particular bispectrum shapes using Table 1.

## 2.2 The mass function in terms of $\mathcal{M}_3$ and the scaling of higher moments

We will assume the non-Gaussian factor in the mass function of Eq.(2.1) takes the following form:

$$\left. \frac{n_{\text{NG}}}{n_{\text{G}}} \right|_{\text{Edgeworth}} \approx 1 + \frac{F_1^{\text{h,f'}}(M)}{F_0'(M)} + \frac{F_2^{\text{h,f'}}(M)}{F_0'(M)} + \dots \quad (2.9)$$

Each term in the series is normalized by the Press-Schechter Gaussian term,  $F_0'(M) = (e^{-\nu_c^2/2}/\sqrt{2\pi})(d\sigma/dM)(\nu_c/\sigma)$ , where  $\nu_c = \delta_c/\sigma$ ,  $\delta_c = 1.686$  is the collapse threshold, and  $\sigma = \sigma(M)$  is the variance in density fluctuations smoothed on the appropriate scale (Eq.(A.4)). Although the first term,  $F_1^{\text{h,f'}}(M)$  or  $F_1^{\text{f'}}(M)$ , is proportional to  $\mathcal{M}_3$  regardless of how the higher moments scale, the exact form of all higher order terms depends on the choice of scaling. For the hierarchical and feeder scaling,  $F_n^{\text{h,f'}}(M)$  and  $F_n^{\text{f'}}(M)$  are given in Eq.(A.14) of the Appendix. Truncating this series after the first term is clearly unphysical since no probability distribution with only a non-zero skewness can be positive everywhere. Although for some objects (low mass, low redshift) this truncation does not cause a significant error, for rarer fluctuations it does. Keeping higher terms in the series is therefore important. How significant these terms are in the context of cluster constraints depends on the mass and redshift of the objects as well as the amplitude and scaling of the non-Gaussianity considered. In Section 5, we show several examples to illustrate how relevant the higher terms are as a function of mass, redshift, skewness and scaling. Although this mass function has been shown to agree reasonably well with simulations, it does not come from a first principles derivation. In Section 5 we also contrast it to the Dalal et al mass function from simulations of the local ansatz [20].

---

<sup>4</sup>The local ansatz was given in Eq.(1.3) with  $f_{\text{NL}}^2 \langle \Phi_{\text{G}}(x)^2 \rangle \ll 1$  to ensure weak non-Gaussianity. The moments generated have the hierarchical scaling with  $\mathcal{I} = f_{\text{NL}}$ . To obtain representative combinatorics for the feeder scaling, we use a scenario where one Gaussian field and one subdominant but highly non-Gaussian field contribute to the inhomogeneities in the gravitational potential:  $\Phi(x) = \phi_{\text{G}} + \sigma_{\text{G}} + \tilde{f}_{\text{NL}}[\sigma_{\text{G}}(x)^2 - \langle \sigma_{\text{G}}(x)^2 \rangle]$ , with  $\tilde{f}_{\text{NL}} \mathcal{P}_{\sigma}^{1/2} \gg 1$ . In that case  $\mathcal{I} = \tilde{f}_{\text{NL}} \mathcal{P}_{\sigma} / \mathcal{P}_{\Phi}^{1/2}$ .

### 3 Two parameter extensions

We also consider the ability of the data to constrain four models characterized by two parameters, chosen to match classes of non-Gaussian inflation models. One very natural extension is to introduce an additional parameter,  $0 \leq q \leq 1$ , so that the hierarchical moments of the non-Gaussian density field behave as

$$\mathcal{M}_n^h = n! 2^{n-3} q \left( \frac{\mathcal{M}_3^h}{6q} \right)^{n-2}. \quad (3.1)$$

Moments of this type arise in scenarios when two fields, one Gaussian and the other with weak local-type non-Gaussianity, contribute to the primordial fluctuations. In that case the total primordial gravitational perturbation is

$$\Phi(x) = \phi_G(x) + \psi_G(x) + \tilde{f}_{\text{NL}}(\psi_G^2 - \langle \psi_G^2 \rangle). \quad (3.2)$$

The ratio of the contribution of (the Gaussian part of) the field  $\psi_G$  to the total power is the parameter  $q$ :

$$q \equiv \frac{\langle \psi_G^2 \rangle}{\langle \phi_G^2 \rangle + \langle \psi_G^2 \rangle} \quad (3.3)$$

and  $\mathcal{I} = q\tilde{f}_{\text{NL}}$ . Since significant local non-Gaussianity only arises in multi-field models, this model is quite plausible. In the standard local ansatz, only one of the terms in Eq.(3.2) contributes to the fluctuations in the gravitational potential, but that field is distinct from the field that sourced inflation. In other words, if  $\phi_G$  represents the fluctuations of the inflaton, the usual local ansatz corresponds to Eq.(3.1) with  $q \rightarrow 1$ . In these inflation models, both fields must be very light compared to the Hubble scale during inflation, with masses  $m_{1,2} \ll H$ .

We can similarly introduce a second parameter into the feeder scaling by defining

$$\mathcal{M}_n^f = q(n-1)! 2^{n-1} \left( \frac{\mathcal{M}_3^f}{8q} \right)^{n/3} \quad (3.4)$$

again with  $0 \leq q \leq 1$ . Such a coefficient appears in the inflation scenario of [31] and in the ‘quasi-single field’ models introduced by [36] where there is an additional heavy field relevant for the fluctuations (with mass very close to the Hubble scale during inflation).

For either choice of scaling for the cumulants, we will take  $\mathcal{M}_3$  and  $q$  as the free parameters. For fixed  $\mathcal{M}_3$ , smaller  $q$  then corresponds to boosting the relative importance of higher order cumulants, while  $q = 1$  corresponds to our one-parameter models. For the scenarios to remain weakly non-Gaussian, we need  $\mathcal{M}_3^{h,f}/q \ll 1$ . If our cosmology is consistent with Gaussian primordial fluctuations, we expect the data to favor  $q \approx 1$  and small  $\mathcal{M}_3$ .

As another extension of the one-parameter models, we add scale-dependence. One advantage of cluster number count constraints is the relatively small scale they probe compared to the CMB. This suggests that clusters may be particularly useful when used in conjunction with the CMB to constrain scale-dependent non-Gaussianity [30]. To that end, we consider a scale-dependent ansatz with hierarchical scaling of the cumulants:

$$\begin{aligned} \mathcal{M}_{3,R} &= 6 \mathcal{I}_{R_0} \left( \frac{R}{R_0} \right)^{n_3}, \\ \mathcal{M}_{n,R} &= n! 2^{n-3} \left( \frac{\mathcal{M}_3}{6} \right)^{n-2}, \end{aligned} \quad (3.5)$$



where  $n_3$  is the additional model parameter. We chose the pivot scale to be  $R_0 = 8h^{-1}$  comoving Mpc. Similarly, we can add scale dependence to the feeder type scenarios:

$$\mathcal{M}_{3,R} = 8 \mathcal{I}_{R_0} \left( \frac{R}{R_0} \right)^{n_3}, \quad (3.6)$$

$$\mathcal{M}_{n,R} = (n-1)! 2^{n-1} \left( \frac{\mathcal{M}_3}{8} \right)^{n/3}.$$

## 4 Application to galaxy cluster data

We now investigate the constraints on these models from current galaxy cluster data. In this work, we employ the data and analysis of [3, 5], to which we refer the reader for complete details.

Briefly, the cluster data set consists of 237 X-ray bright clusters detected in the ROSAT All-Sky Survey [4] and compiled in the BCS, REFLEX or bright MACS catalogs [37–39] satisfying conservative luminosity/flux thresholds such that the physical properties and selection function of the final sample are well understood.<sup>5</sup> These detections span the redshift range  $0 < z < 0.5$ ; in Figure 1 their distribution in redshift and mass is compared to SZ detections that have been used in previous non-Gaussianity constraints. In addition to survey X-ray flux measurements and spectroscopic redshifts, archival ROSAT or *Chandra* data for 94 of the clusters were used to obtain mass proxies in the form of temperature and gas mass measurements. Total mass measurements from hydrostatic analysis of *Chandra* data for 42 relaxed clusters were also incorporated [40].<sup>6</sup> Together these data provide the ability to calibrate the scaling and intrinsic scatter with mass of cluster X-ray luminosity, gas temperature and gas mass. This information is critical to accounting for survey selection effects and provides improved cosmological constraints generally compared with a pure self-calibration approach. Full details of our scaling relation model and the data used to constrain it are found in [3, 5].

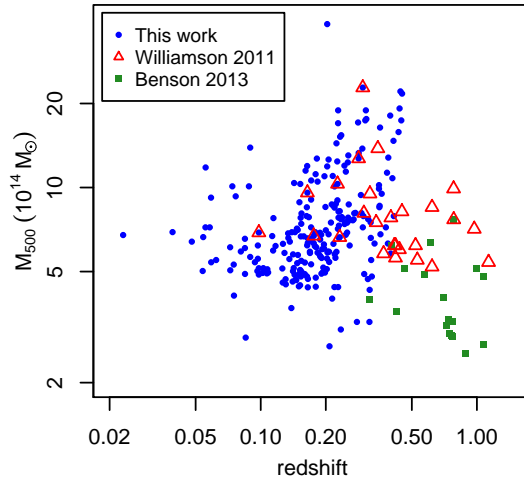
Through a technique detailed in [40], measurements of the ratio of gas mass to total mass for the 42 relaxed clusters above additionally and independently provide constraints on the mean cosmic matter density and the expansion of the Universe, which we also take advantage of here. We additionally incorporate seven-year WMAP constraints in some of the results below, using the publicly released WMAP data and analysis codes [35]. We do not model the effects of non-Gaussianity on the CMB power spectrum in this analysis, which is justified since even the current constraints limit any shift due to non-Gaussianity to be small; consequently, the WMAP data effectively only provide additional constraining power on the standard set of cosmological parameters, particularly  $\sigma_8$ .

Following the method outlined in Section 2, we evaluate the non-Gaussian mass function as the product of a Gaussian mass function and a series whose terms depend on the model under consideration (see Eq.(2.1) and Eq.(A.13)). As in [3, 5], the Gaussian mass function we use is the simulation-calibrated fit of [33], with the cosmology-dependent correspondence of halo mass and  $\sigma(M)$  provided by CAMB.<sup>7</sup> Note that, independent of any non-Gaussianity, our treatment includes an allowance for systematic uncertainty in the normalization, shape,

<sup>5</sup>Since the original work, Abell 689 has been removed from the data set, since *Chandra* observations reveal its X-ray emission to be dominated by a point source rather than the intracluster medium.

<sup>6</sup>Note that this sample of 42 only partly overlaps the larger cosmology sample, a fact which is accounted for in the data analysis.

<sup>7</sup><http://www.camb.info>

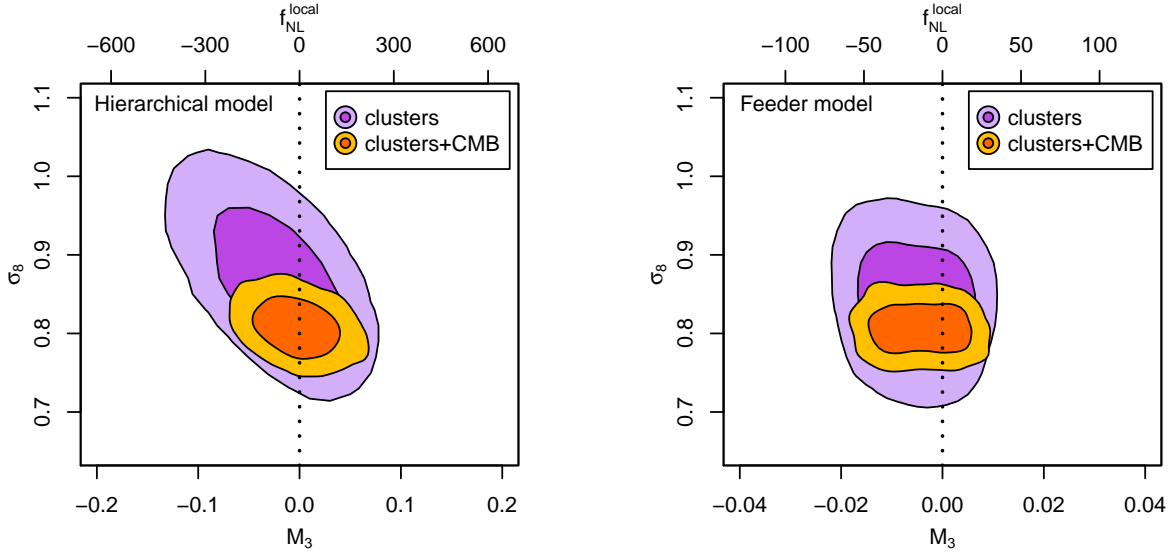


**Figure 1.** For illustrative purposes, the mass and redshift distribution of our cluster data [5] are compared with South Pole Telescope detected clusters used to obtain constraints on  $f_{\text{NL}}^{\text{local}}$  by Williamson et al. [16] and Benson et al. [15]. The mass values assume a flat concordance model with  $h = 0.7$  and  $\Omega_m = 0.3$ .

and evolution of the Gaussian mass function in the form of a multidimensional prior on the relevant parameters [3]. This prior translates to a 10% uncertainty on the Gaussian mass function at the typical mass of our clusters ( $10^{15} M_\odot$ ). We evaluate the non-Gaussian series given by Eq.(A.13) and Eq.(A.14), keeping at most 16 terms for the hierarchical model and 17 for the feeder. This is a large enough number of terms that the series expansion of the PDF (estimated by the relative size of the first ignored term) is accurate to at least 20% for the highest redshift, most massive clusters in the sample and  $\mathcal{M}_3 < 0.04$  (0.025) for hierarchical (feeder) scaling. However, for the low values of  $\mathcal{M}_3$  that the data prefers, a considerably smaller number of terms (around 5) is sufficient for the PDF to be accurate to a few percent in the range of interest, and there is not much difference in the mass function as the number of terms is changed. Furthermore, for larger levels of non-Gaussianity (outside the observational bounds, but still weakly non-Gaussian) the series can be badly behaved if too many terms are kept. See [32] for more detail.

Constraints on non-Gaussianity from clusters are limited by the precision with which cluster masses can be estimated, both individually (i.e. constraining the masses of the most rare objects) and statistically (constraining the mass function of the population). The former is straightforwardly related to the limitations of individual cluster measurements, while the latter also depends on the scaling and intrinsic scatter with mass of the survey observable used to define the cluster sample, as well as the ability of the data to constrain those nuisance parameters. In the present work, both our estimates of individual masses and the overall cluster mass scale include a systematic error budget of  $\sim 15\%$ ; these reflect allowances for a variety of uncertainties such as instrument calibration and departures from hydrostatic equilibrium [5, 40]. Including all these allowances, our data ultimately provide a 10% constraint on the normalization of the X-ray luminosity–mass relation, with the best fitting intrinsic scatter of that relation measured to be  $(43 \pm 4)\%$  [5].

When analyzing cluster data, we vary the mean baryon and matter densities ( $\Omega_b$  and  $\Omega_m$ ), the Hubble parameter ( $h$ ) and the amplitude of the matter power spectrum ( $\sigma_8$ ), in addition to the parameters describing non-Gaussianity and cluster scaling relations, and a

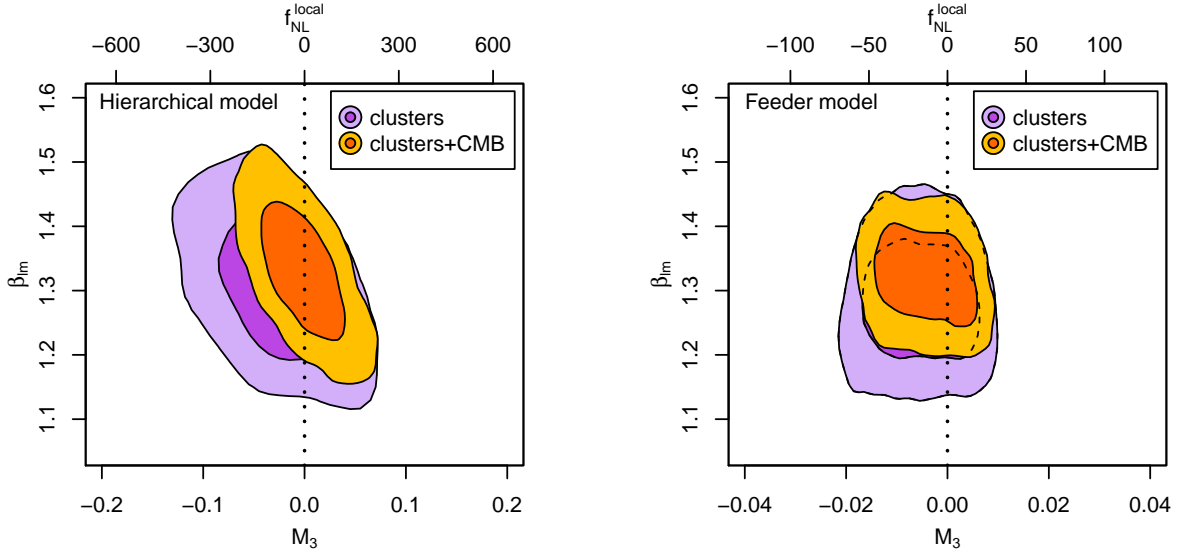


**Figure 2.** Marginalized joint constraints on  $\mathcal{M}_3$  (or the corresponding  $f_{\text{NL}}^{\text{local}}$ ) and  $\sigma_8$  at 68.3% and 95.4% confidence levels from cluster and clusters+CMB data, for single-parameter non-Gaussian models. Only CMB power spectra (not bispectra) are used in the combination, tightening constraints on the standard set of cosmological parameters, but providing no additional, direct constraining power on non-Gaussianity. The constraints are consistent with Gaussianity in all cases, but their strength and character depend on whether the hierarchical or feeder scaling (left and right panels) are used.

number of nuisance parameters accounting for various systematic uncertainties (see [3, 40]). Priors on the Hubble parameter,  $h = 0.742 \pm 0.036$  [41], and on the baryon density from Big Bang Nucleosynthesis (BBN),  $\Omega_b h^2 = 0.0214 \pm 0.002$  [42], are included. When including WMAP data, we additionally marginalize over the optical depth to reionization, the spectral index of scalar fluctuations and its running with wavenumber, and the amplitude of small-scale CMB fluctuations due to the Sunyaev-Zel’dovich effect; the Hubble parameter and BBN priors are not used in this case. The full suite of our results are shown in Tables 2 and 3.

Marginalized joint constraints on  $\mathcal{M}_3$  and  $\sigma_8$  are shown in Figure 2 for the single-parameter non-Gaussian model, using both hierarchical and feeder scalings. In both cases, the principal degeneracy of  $\mathcal{M}_3$  is with  $\sigma_8$  when only cluster data are used. This is intuitive, since  $\sigma_8$  determines the relative rarity of massive clusters even in the purely Gaussian case. With the addition of CMB data,  $\sigma_8$  is independently tightly constrained, and for both scalings the primary degeneracy of  $\mathcal{M}_3$  is instead with the slope of the cluster X-ray luminosity–mass relation,  $\beta_{\ell m}$ , as shown in Figure 3. This reflects the dependence of the number of *detections* of massive clusters in an observable-limited survey on the relevant scaling relation. Compared with the hierarchical scaling, the degeneracy between  $\mathcal{M}_3$  and other model parameters is generically weaker in the feeder model. This is due to the relatively larger high-order non-Gaussian moments generated by the feeder scaling, whose effect is less easily mimicked by changing other parameters (Figure 6; see also Figure 1 of [28]).

We also consider two-parameter non-Gaussian models for the two scalings, including either the  $q$  or  $n_3$  parameters introduced in Section 3. In neither case is the additional parameter well constrained by the data, although the impacts on the  $\mathcal{M}_3$  constraints are relatively minor (Table 2). However, we find that values  $q \approx 0$ , corresponding to having strongly non-Gaussian fluctuations regardless of the amplitude of  $\mathcal{M}_3$ , are disfavored (see



**Figure 3.** As in Figure 2, but showing constraints on  $\mathcal{M}_3$  and  $\beta_{\ell m}$ , the power-law slope of the cluster X-ray luminosity–mass relation.

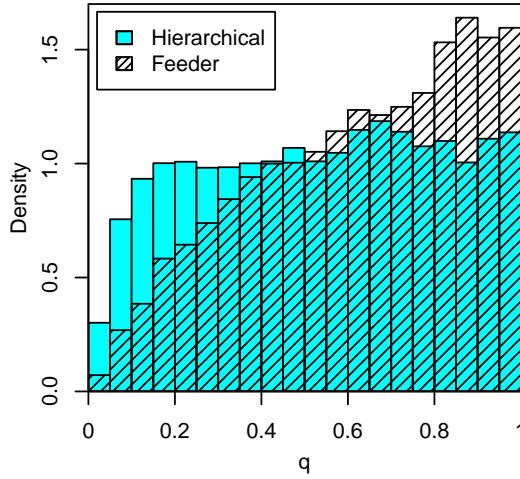
**Table 2.** Marginalized best fitting values and 68.3% confidence intervals on the most interesting fit parameters (see text), using cluster (CL) and CL+CMB data, for various models (choices of the scaling and the number of free parameters). ‘skew-only’ refers to truncating the series in Eq.(A.13) after the second term, in which case there is no distinction between hierarchical (h) and feeder (f) scalings. For the scale-dependent models, the amplitude of the skewness is reported at the pivot point. <sup>a</sup>A small local maximum in probability in the range  $[-33, -28]$  is also formally included in this maximum likelihood confidence region (see also Figure 5).

Model	Data	$10^3 \mathcal{M}_3$	$\sigma_8$	$\beta_{\ell m}$
h	CL	$-23^{+40}_{-35}$	$0.86^{+0.07}_{-0.06}$	$1.29^{+0.08}_{-0.07}$
h	CL ( $z < 0.3$ )	$113^{+70}_{-65}$	$0.76^{+0.08}_{-0.05}$	$1.28^{+0.08}_{-0.10}$
h	CL+CMB	$-1^{+24}_{-28}$	$0.81^{+0.02}_{-0.03}$	$1.33^{+0.07}_{-0.08}$
h+q	CL+CMB	$-7^{+28}_{-24}$	$0.81^{+0.03}_{-0.02}$	$1.34^{+0.07}_{-0.08}$
h+n <sub>3</sub>	CL+CMB	$3^{+20}_{-30}$	$0.81^{+0.02}_{-0.03}$	$1.31^{+0.09}_{-0.06}$
f	CL	$-9^{+11}_{-4}$	$0.83^{+0.06}_{-0.04}$	$1.27^{+0.07}_{-0.06}$
f	CL ( $z < 0.3$ )	$22^{+11}_{-19}{}^a$	$0.84^{+0.06}_{-0.06}$	$1.32^{+0.08}_{-0.09}$
f	CL+CMB	$-4^{+7}_{-7}$	$0.81^{+0.02}_{-0.02}$	$1.32^{+0.06}_{-0.05}$
f+q	CL+CMB	$-1^{+4}_{-5}$	$0.81^{+0.02}_{-0.02}$	$1.32^{+0.06}_{-0.05}$
f+n <sub>3</sub>	CL+CMB	$-5^{+8}_{-6}$	$0.80^{+0.02}_{-0.02}$	$1.32^{+0.07}_{-0.05}$
skew-only	CL	$-9^{+164}_{-24}$	$0.81^{+0.05}_{-0.09}$	$1.26^{+0.07}_{-0.06}$
skew-only	CL+CMB	$-3^{+72}_{-20}$	$0.80^{+0.03}_{-0.03}$	$1.28^{+0.08}_{-0.04}$

Figure 4). The clusters+CMB data respectively provide 95.4% confidence lower limits of

**Table 3.** The constraints on the skewness can be converted to constraints on the amplitude of any bispectrum. The shape of the bispectrum is independent of the scaling, although the usual local ansatz corresponds to a local-shape bispectrum with hierarchical moments.

Scaling	Data	Local Bispectrum	Equil. Bispectrum	Orthog. Bispectrum
h	CL	$-73^{+129}_{-113}$	$-271^{+482}_{-422}$	$346^{+538}_{-615}$
h	CL+CMB	$-3^{+78}_{-91}$	$-12^{+289}_{-338}$	$15^{+430}_{-369}$
f	CL	$-28^{+35}_{-13}$	$-106^{+134}_{-48}$	$130^{+60}_{-164}$
f	CL+CMB	$-14^{+22}_{-21}$	$-52^{+85}_{-79}$	$63^{+97}_{-104}$
skew-only	CL	$-29^{+532}_{-78}$	$-105^{+1916}_{-280}$	$146^{+389}_{-2658}$
skew-only	CL+CMB	$-9^{+234}_{-65}$	$-35^{+841}_{-234}$	$48^{+324}_{-1167}$

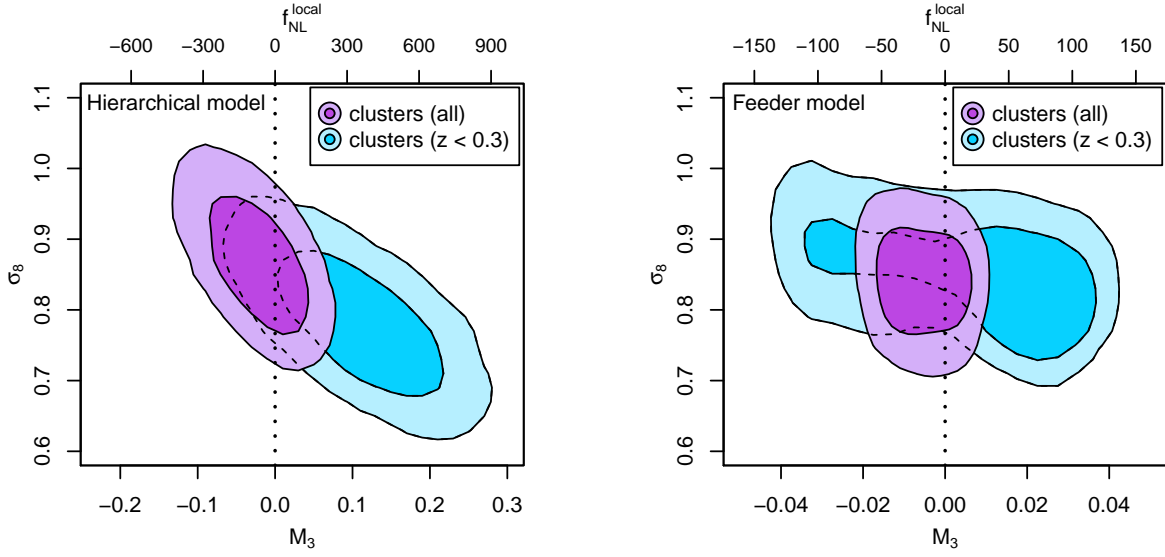


**Figure 4.** Posterior probability density of the  $q$  parameter for hierarchical and feeder scalings. In our parametrization, small values of  $q$  boost higher order non-Gaussian moments relative to  $\mathcal{M}_3$  and  $q = 1$  corresponds to the one-parameter models with either scaling. The clusters+CMB data disfavor very small values of  $q$ , consistent with the overall preference for Gaussian initial fluctuations, although only the feeder scaling shows a clear (if modest) preference for  $q = 1$ . At 95.4% confidence, the lower limits on  $q$  are respectively  $q > 0.10$  and  $q > 0.18$  for the hierarchical and feeder scenarios.

$q > 0.10$  and  $q > 0.18$  with the hierarchical and feeder scalings, assuming a uniform prior from zero to one.

For ease of comparison to the literature (Section 5), we also obtained constraints on  $\mathcal{M}_3$  keeping only the first term in the non-Gaussian mass function, proportional to the skewness. As shown in Table 2, the resulting error bars are larger than when we include more terms in the series, reflecting the fact that those terms generically increase the deviation of the mass function from the Gaussian one at higher masses and higher redshifts (again, see Figure 6). For example, keeping only the first term, our constraints from cluster and CMB data correspond to  $f_{\text{NL}}^{\text{local}} = -9^{+234}_{-78}$  (Table 3). When including all relevant terms we find  $f_{\text{NL}}^{\text{local}} = -3^{+78}_{-91}$  for the hierarchical scaling and  $f_{\text{NL}}^{\text{local}} = -14^{+22}_{-21}$  for the feeder scaling.

In light of the favorable comparison of our results to those obtained previously (see Sec-



**Figure 5.** As in Figure 2, but comparing constraints obtained from the full cluster data set with those from only clusters at redshifts  $z < 0.3$ .

tion 5), we briefly investigate what characteristics of our data set influence the results. There is a practical limitation, however, since any attempt to reduce the overall size of the sample, the number of clusters with mass estimates from follow-up data, or the mass/redshift ranges covered, necessarily impacts constraints on the full set of cosmological and scaling relation parameters. Consequently, we confine ourselves to a single, limited, but informative comparison by asking how our constraints change when data at  $z \geq 0.3$  are excluded. In detail, this low-redshift sample contains 203 clusters, of which 61 have follow-up data, compared to 237 and 94 for the full data set. As shown in Figure 5 for single-parameter non-Gaussian models using the full hierarchical and feeder scalings, the constraining power of this low-redshift data set is significantly reduced. Results for the low-redshift clusters only are shown in Table 2.

## 5 Comparison with the literature

Previous forecasts for constraints on non-Gaussianity from cluster counts have been done by Pillepich et al. [7] for the eROSITA X-ray mission, Sartoris et al. [43] for future X-ray surveys resembling the Wide Field X-ray Telescope concept, Oguri [44] for a variety of future optical surveys, Cunha et al. [6] for optically selected clusters in the Dark Energy Survey (DES), and Mak and Pierpaoli [8] for future surveys using the Sunyaev-Zel’dovich effect. There have been three previous cluster constraints on non-Gaussianity: two based on clusters detected in the SPT survey, by Benson et al. [15], who find  $f_{\text{NL}}^{\text{local}} = -192 \pm 310$ , and Williamson et al. [16], who report  $f_{\text{NL}}^{\text{local}} = 20 \pm 450$ ; and one based on the SDSS maxBCG cluster catalogue, by Mana et al. [17], who have  $f_{\text{NL}}^{\text{local}} = 282 \pm 317$ .<sup>8</sup>

The existing forecasts and constraints use a variety of prescriptions for the non-Gaussian mass function (listed in Table 4). These mass functions are given in Eq.(A.19) in the Appendix and are plotted in Figure 6. The levels of non-Gaussianity shown are  $\mathcal{M}_3 = 0.009$

<sup>8</sup>This result corresponds to their analysis of only cluster number counts, without including either the cluster power spectrum or CMB data.



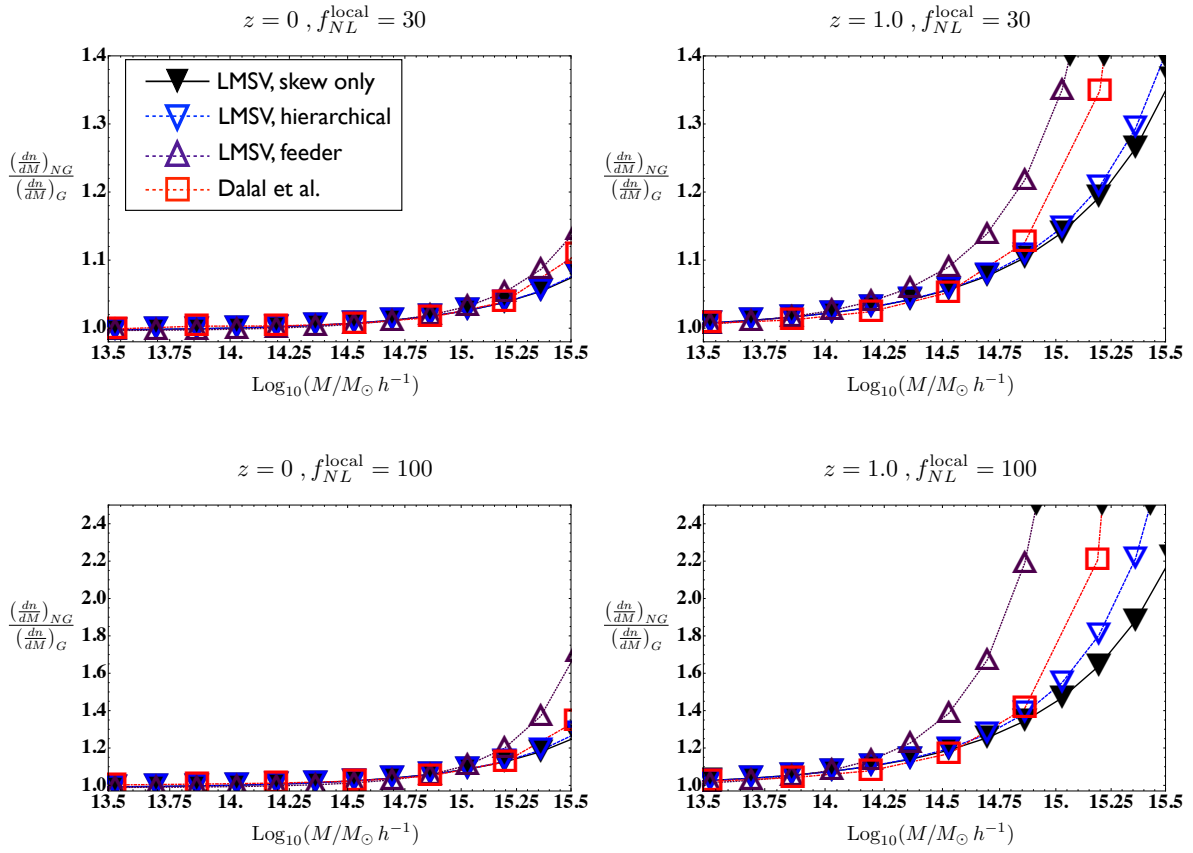
and  $\mathcal{M}_3 = 0.031$ , which correspond to the local model bispectrum with  $f_{\text{NL}}^{\text{local}} = 30$  and 100, respectively. Notice that, below about  $10^{15} M_{\odot}$ , the Dalal et al. mass function [20] deviates a little less from the Gaussian for a given value of  $f_{\text{NL}}$  than the LoVerde et al. mass function [30] (hereafter LMSV). However, some authors have found the LMSV mass function agrees better with simulation results if a reduced collapse threshold,  $\delta_c \sim 1.5$ , is used. If that adjustment is made, the Dalal et al mass function would deviate more from the Gaussian than LMSV; see [45] for a comparison of all these cases. Since the Dalal et al. mass function was calibrated on simulations of the local ansatz, in principle it should include information about higher moments. This technique, though, has only been tried against one set of simulations and only for non-Gaussianity of the local type. A more precisely calibrated, more general non-Gaussian mass function will be important for any future analysis of non-Gaussianity with clusters.

**Table 4.** Gaussian mass functions and non-Gaussian extensions used in the literature. The non-Gaussian mass functions are either the first order semi-analytic expression from LoVerde et al. [30] (LMSV) or the mass function calibrated on N-body simulations of the local ansatz by Dalal et al. [20]. All non-Gaussian mass functions also make use of a Gaussian mass function such as those fit by Sheth and Tormen [46], Warren et al. [47], Jenkins et al. [48] or Tinker et al. [33, 49].

Author	Mass Function used
Benson [15]	Jenkins + $f_{\text{Dalal}}$
Cunha [6]	Jenkins + $f_{\text{Dalal}}$
Mak [8]	Tinker + $f_{\text{LMSV,skew only}}$
Mana [17]	Tinker + $f_{\text{LMSV,skew only}}$
Oguri [44]	Warren + $f_{\text{LMSV,skew only}}$
Pillepich [7]	Tinker + $f_{\text{LMSV,skew only}}$
Sartoris [43]	Sheth-Tormen + $f_{\text{LMSV,skew only}}$
Williamson [16]	Jenkins + $f_{\text{Dalal}}$
This work	Tinker + $f_{\text{LMSV,many terms}}$

Apart from the non-Gaussian mass function, these forecasts and analyses differ from one another and from ours in two principal ways: the form and complexity assumed for the mass–observable relation and its intrinsic scatter, and priors on the associated parameters. The most pessimistic forecasts in the literature find marginalized one sigma errors on  $f_{\text{NL}}^{\text{local}}$  around  $\mathcal{O}(10^3)$  (e.g. some cases analyzed in [6, 7]). Those results assume that the scaling relations will be constrained solely through self-calibration [50] rather than with estimates of cluster masses, which can significantly boost the constraining power [51]. In addition, some forecasts assume significant photometric redshift errors [7]. As outlined in Section 4, all of the clusters in our sample have spectroscopic redshifts and for nearly half we also have follow-up X-ray data that significantly improve the mass determinations.

Among the SPT results, Benson et al. use a smaller area of the survey than Williamson et al., but have an improved mass calibration and extend their sample to lower SZ detection significance (i.e. lower mass). In comparison, our cluster data set is significantly larger than either of the SPT cluster samples, contains more massive clusters (although at lower redshifts), has a larger intrinsic scatter in the mass–observable relation (although the parameters of the scaling relation are better constrained), and uses a more straightforward mass calibra-



**Figure 6.** Comparing mass functions. Black (filled point down triangles): LMSV with one term. Blue (open point down triangles): LMSV with hierarchical scaling, keeping 4 terms. Purple (open point up triangles): LMSV with feeder scaling, keeping seven terms. (The number of terms was chosen to give good behavior up to  $f_{\text{NL}} = 300$  at redshift 1.) Red (squares): Dalal et al. + Tinker et al. Note the difference in scale on the vertical axes in the top row compared to the bottom row.

tion (i.e. directly incorporating X-ray mass measurements rather than calibrating mass via simulation priors or priors based on external X-ray data). The masses and redshifts of these three cluster samples are compared in Figure 1.

Both of the SPT studies used CMB data to constrain the amplitude of the Gaussian power spectrum, as well as other cosmological parameters, as we did in Section 4. Both also used the Dalal et al. non-Gaussian mass function, so it is not immediately clear which of our constraints is most comparable to theirs. However, comparing the various mass functions for the masses and redshifts of the SPT clusters and non-Gaussianity of magnitude  $f_{\text{NL}}^{\text{local}} \sim \mathcal{O}(100)$ , the better comparison seems to be with our skew-only results (Figure 6). Our clusters+CMB constraint for the skew-only model corresponds to  $f_{\text{NL}}^{\text{local}} = -9_{-65}^{+234}$ , roughly a factor of two tighter than the constraints of Benson et al., but less of an improvement than a straightforward dependence on the cluster sample size (237 versus 18) would imply. The precision of the mass calibration used in the two works is similar, and most likely limits the improvement we see from the larger data set. The greater constraining power of the Benson et al. analysis (18 clusters) versus that of Williamson et al. (26 clusters) further underscores

the impact of providing additional data to improve the cluster mass calibration. On the other hand, the significant improvement in the constraining power of our full data set compared with only the low-redshift portion (see Figure 5) indicates that our current constraints are still not entirely systematically limited.

Compared to both our data and the SPT samples, the maxBCG catalogue used by Mana et al. [17] contains a very large number (13 823) of mostly less massive clusters at low redshifts ( $0.1 < z < 0.3$ ) [29]. Their mass calibration is accomplished through a stacked weak gravitational lensing analysis, which constrains the mean optical richness–mass relation but does not provide mass estimates for individual clusters [52, 53]. This feature, as well as the lack of high mass and high redshift clusters, presumably limits their constraining power. We note, however, that the large range in mass probed allows them to achieve from cluster counts *alone* constraints comparable to those of Benson et al. from clusters+CMB. Our skew-only, clusters-only constraint ( $f_{\text{NL}}^{\text{local}} = -29_{-78}^{+532}$ ) is comparable to that of Mana et al.

There are some existing constraints on primordial non-Gaussianity beyond the skewness that are complementary to those we find here. For example, the CMB data from the WMAP satellite has been analyzed by [10, 54]. The *Planck* data has not yet been exhaustively analyzed for evidence of a trispectrum (a four-point correlation function), but the collaboration has reported the strongest bound to date for one of the simplest momentum-space trispectrum shapes. This shape is part of the local family of non-Gaussianity (it is the four point function generated by Eq.(1.3)), but its amplitude can be constrained independently of the amplitude of the three-point function ( $f_{\text{NL}}^{\text{local}}$ ) and is typically called  $\tau_{\text{NL}}$ . *Planck* data constrains the amplitude of this shape to be  $\tau_{\text{NL}} < 2800$  [10], which is much weaker than the constraint implied by *Planck*’s tight limits on  $f_{\text{NL}}^{\text{local}}$  for a single parameter model with hierarchical scaling. However, the weakness of this constraint is likely to be due in large part to the remaining foreground and systematic effects present in and yet to be removed from the current analysis. For comparison, our central value in the hierarchical case,  $\mathcal{M}_3 = -0.001$  corresponds to a model with  $\tau_{\text{NL}} \approx 9$  (using the fitting function of [55] at  $10^{14} h^{-1} M_{\odot}$ ). Our central value for the feeder scaling, which is much more non-Gaussian, corresponds to  $\tau_{\text{NL}} \approx 12\,300$ . By introducing the parameter  $q$ , we also tested models where the relationship between the skewness and the kurtosis was relaxed, although the rest of the series was still determined in terms of the two parameters. The constraints on non-Gaussianity from *Planck* data using Minkowski functionals do incorporate information from the series of cumulants in a way similar to cluster number counts. These could be more directly compared to our results with some additional analysis.

Recently Giannantonio et al [14] reported a constraint on non-Gaussianity from a clustering analysis of a broad ensemble of multiwavelength galaxy surveys. In that analysis,  $f_{\text{NL}}^{\text{local}}$  is degenerate with a trispectrum parameter  $g_{\text{NL}}$  (the amplitude of the bispectrum generated by adding a term proportional to  $\Phi_G^3$  to the local ansatz). For  $f_{\text{NL}}^{\text{local}} = 0$  they report  $-4.5 \times 10^5 < g_{\text{NL}} < 1.6 \times 10^5$  (95% C.L.) while if  $f_{\text{NL}}^{\text{local}}$  is nonzero and positive they find that  $g_{\text{NL}}$  can take somewhat more negative values. Our results, using the hierarchical scaling appropriate for weakly non-Gaussian local models, are comparable. Again using a fitting formula from [55] at  $10^{14} h^{-1} M_{\odot}$ , our largest allowed magnitude for the kurtosis ( $|\mathcal{M}_4| \sim 5 \times 10^{-3}$  for the one-parameter model with approximate 95% bounds  $|\mathcal{M}_3| \lesssim 60 \times 10^{-3}$ ) corresponds to  $|g_{\text{NL}}| \lesssim 9 \times 10^4$ . Allowing the two-parameter model with  $q \sim 0.1$  relaxes this constraint to  $|g_{\text{NL}}| \lesssim 9 \times 10^5$ . If the same trispectrum shape is considered with feeder scaling for the rest of the moments, our results imply  $|g_{\text{NL}}| \lesssim 3 \times 10^5$  for the single parameter model and  $|g_{\text{NL}}| \lesssim 6 \times 10^5$  with  $q = 0.18$ .

## 6 Conclusions

We have used the mass and redshift distribution of X-ray bright clusters to find that the primordial inhomogeneities in the gravitational potential can be consistently described by a Gaussian distribution to an accuracy of about one part in  $10^3$  at scales around  $0.1\text{--}0.5 h \text{ Mpc}^{-1}$ . Our constraints apply to any model of weak non-Gaussianity that has a sufficiently regular ordering in the cumulants to be modeled by one of the expressions in Section 2 or Section 3 (Equations (2.7), (2.8), (3.1), (3.4), (3.5) or (3.6)). In particular, for non-Gaussianity described by the local ansatz (Eq.(1.3), which defines all cumulants for the model), we find  $f_{\text{NL}}^{\text{local}} = -3_{-91}^{+78}$  from cluster data combined with the WMAP 7-year data.

Our analysis differs from previous work in that we include higher order cumulants in the non-Gaussian mass function. This allows us to differentiate constraints on two different one-parameter models for the non-Gaussianity, characterized by the relevance of moments beyond the skewness. We also tested the sensitivity to several two-parameter models, but found that current data are not very sensitive to this extra level of detail. Our full set of results can be found in Table 2. Table 3 shows those results interpreted in terms of several popular models for the bispectrum.

The *Planck* satellite data recently led to very tight constraints on any non-Gaussianity on a significant range of scales, but those bounds are still about two orders of magnitude above the minimal levels predicted by slow-roll inflation. Some combination of lower-redshift probes will be required to explore the rest of that parameter space. This is an extremely worthwhile pursuit, since higher order correlations (or their absence) is our only route to learn more about the primordial era. The previously unexplored sensitivity of cluster data to higher order moments of non-Gaussianity shown by our results argues for revisiting the forecasts for future cluster surveys. Further bounds on non-Gaussianity remain as theoretically interesting as more precise measurements of the dark energy equation of state. Since the next stage of research on the primordial era will be dominated by large scale structure observations, it is crucial to understand the full potential of complementary information from cluster number counts together with CMB and large scale structure constraints on the bispectrum and trispectrum, and the halo bias.

There are several ways our analysis could be extended. First, in order to consistently treat as large a family of non-Gaussian models as possible, we have used a simple, semi-analytic form for the non-Gaussian mass function. It would be preferable to use non-Gaussian mass functions calibrated on simulations, but those results do not yet exist for a sufficient variety of scenarios. However, simulations of a two-field local model that spans between the hierarchical scenario (with the extra parameter  $q$ ) and the feeder behavior are in progress [56]. Once this simulation work is complete, we could revisit our analysis using the Dalal et al. mass function for the local ansatz together with a comparable expression for models whose higher moments are relatively more important.

The cluster data sets that have been used to constrain non-Gaussianity so far differ qualitatively in several respects. Nevertheless, empirical comparison of their constraining power underscores that the precision of the overall mass calibration, the availability of mass estimates for individual clusters, and the mass and redshift ranges probed all have an important role. In the very near term, combining the available survey data as well as gravitational lensing data (e.g. [57, 58]) in a multi-wavelength analysis has significant potential.

A variety of upcoming survey data could also improve on the non-Gaussianity constraints in complementary ways. Data from, for example, the optical-wavelength Dark En-

ergy Survey [59] and the upcoming eROSITA all-sky X-ray survey [60] will extend the mass range of the cluster samples downward over a wide redshift range, and enable clustering analyses. Continued SZ surveys will be crucial for detecting the most massive, high redshift clusters which are most sensitive to non-Gaussianity. In addition, targeted X-ray follow-up of individual clusters will allow the more precise mass measurements that significantly aid the overall statistical power of cluster samples. All of this data is being collected to study important outstanding problems in cosmology, especially dark energy and neutrino mass, but is also well suited to further tests of non-Gaussianity to study inflation. Constraints on the primordial fluctuations from cluster counts and clustering provide a complementary cross-check to the CMB and to other large scale structure probes, and will continue to be an important tool for cosmology.

## Acknowledgments

S.S. thanks Adrienne Erickcek for carefully checking the properties of higher order terms in the Petrov expansion. We also thank Neal Dalal, Dragan Huterer, Marilena LoVerde and Christian Reichardt for useful comments on a draft of this article. S. Shandera is supported by the Eberly Research Funds of The Pennsylvania State University. The Institute for Gravitation and the Cosmos is supported by the Eberly College of Science and the Office of the Senior Vice President for Research at the Pennsylvania State University. A. Mantz is supported by NSF grant AST-0838187. D. Rapetti acknowledges support from the DARK Fellowship program at the Dark Cosmology Centre, which is funded by the Danish National Research Foundation. Calculations in this work used the Coma and Orange compute clusters at the SLAC National Accelerator Laboratory. This work was supported in part by the U.S. Department of Energy under contract number DE-AC02-76SF00515.

## A Non-Gaussian mass functions

To model the effect on non-Gaussianity on the cluster mass function, we use an extended Press-Schechter approach [61], following previous work in [30, 55, 62–64]. The non-Gaussian probability distribution of (normalized) density fluctuations smoothed on a scale  $R$  (associated to halos of mass  $M$  by  $R = ((3/4\pi\rho)M)^{1/3}$ ) is  $P(\nu, M)$ . Here  $\nu = \delta/\sigma_R$  (with  $\sigma_R$  being the smoothed variance). The fraction of volume in collapsed objects (halos) is

$$F(M) = 2 \int_{\delta_c/\sigma_R}^{\infty} d\nu P(\nu, M) \quad (\text{A.1})$$

with  $\delta_c$  the threshold for collapse, and where we have included the Press-Schechter factor of 2 in front of the integral. Then the number density  $dn(M)/dM$  of halos with masses between  $(M, M + dM)$  is

$$\frac{dn}{dM}(M, z) = -2 \frac{\bar{\rho}}{M} \frac{dF}{dM}, \quad (\text{A.2})$$

where  $\bar{\rho} = \Omega_m \rho_{\text{crit}}$  is the average (comoving) matter density. The smoothed density field is given by

$$\delta_R(z) = \int \frac{d^3k}{(2\pi)^3} W_R(k) \delta(\vec{k}, z) \quad (\text{A.3})$$

where  $W_R(k)$  is the Fourier transform of a window function, which we take to be a top-hat in real space. The smoothed variance is

$$\sigma^2(M, z) = \frac{1}{2\pi^2} \int dk P_{\text{lin}}(k, z) W^2(k, M) k^2. \quad (\text{A.4})$$

The matter perturbations  $\delta$  at redshift  $z$  are related to the perturbations in the early matter era potential  $\Phi$  by

$$\delta(\vec{k}, z) = M(k, z) \Phi(\vec{k}) \quad (\text{A.5})$$

$$M(k, z) = \frac{2}{3} \frac{1}{\Omega_m} \frac{c^2}{H_0^2} D(z) \frac{g(0)}{g(\infty)} T(k) k^2.$$

Here  $\Omega_m$  is the matter density relative to critical,  $H_0$  is the Hubble constant,  $D(z)$  is the linear growth function at redshift  $z$  normalized to one today, and the growth suppression factor is  $g(z=0)/g(z=\infty) \simeq 0.76$  in the best-fit  $\Lambda$ CDM model. The variance of density fluctuations at redshift  $z$  smoothed on a scale  $R$  associated to mass  $M$  is  $\sigma^2(M, z)$ , defined by

$$\sigma^2(M, z) = \int_0^\infty \frac{dk}{k} W_R(k)^2 M(k, z)^2 \Delta_\Phi^2(k), \quad (\text{A.6})$$

where  $\Delta_\Phi^2(k)$  is the amplitude of fluctuations, related to the power spectrum by  $P_\Phi(k) = (2\pi^2/k^3) \Delta_\Phi^2(k) (k/k_0)^{n_s-1}$ .

Petrov [65, 66] (references are the English translations) developed an asymptotic expansion for non-Gaussian PDFs, which is a generalization of the Edgeworth expansion. In terms of the dimensionless smoothed moments  $\mathcal{M}_{n,R}$ , this is

$$P(\nu) d\nu = \frac{d\nu}{\sqrt{2\pi}} e^{-\nu^2/2} \left\{ 1 + \sum_{s=1}^\infty \sum_{\{k_m\}} H_{s+2r}(\nu) \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right\}. \quad (\text{A.7})$$

Here  $H_n(\nu)$  are Hermite polynomials defined by  $H_n(\nu) = (-1)^n e^{\nu^2/2} \frac{d^n}{d\nu^n} e^{-\nu^2/2}$  and  $r = k_1 + k_2 + \dots + k_n$  where the set  $\{k_m\}$  is built of all non-negative integer solutions of

$$k_1 + 2k_2 + \dots + nk_n = s. \quad (\text{A.8})$$

We use the two scalings, hierarchical and feeder, from Eq.(2.7) or Eq.(2.8), to organize the terms in this series. Then we can write

$$F(M) = F_0(M) + F_1(M) + F_2(M) + \dots \quad (\text{A.9})$$

where

$$F_0(M) = \frac{1}{2} \text{Erfc} \left( \frac{\nu_c}{\sqrt{2}} \right) \quad (\text{A.10})$$

and the rest of the series is ordered according to the scaling:

$$F^h(M) = \frac{1}{2} \text{erfc} \left( \frac{\nu_c}{\sqrt{2}} \right) + \frac{e^{-\nu_c^2/2}}{\sqrt{2\pi}} \sum_{s=1}^\infty \sum_{\{k_m\}_h} H_{s+2r-1}(\nu_c) \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \quad (\text{A.11})$$

$$F^f(M) = \frac{1}{2} \text{erfc} \left( \frac{\nu_c}{\sqrt{2}} \right) + \frac{e^{-\nu_c^2/2}}{\sqrt{2\pi}} \sum_{s=1}^\infty \sum_{\{k_m\}_f} H_{s+1}(\nu_c) \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m}.$$



The sets  $\{k_m\}_h$  are again non-negative integer solutions to  $k_1 + 2k_2 + \dots + sk_s = s$  and  $r = k_1 + k_2 + \dots + k_m$ , but the sets  $\{k_m\}_f$  are solutions to  $3k_1 + 4k_2 + \dots + (s+2)k_s = s+2$ . Now for either scaling, truncating the series at some finite  $s$  in the sums above keeps all terms up to the same order in  $\mathcal{M}_3$ :  $\mathcal{M}_3^s$  for hierarchical scalings and  $\mathcal{M}_3^{s/3}$  for feeder scalings.

To write the mass function we will need derivatives of all the terms in the expansion with respect to mass (or smoothing scale). In general, the derivatives can be found using the relationship for the Hermite polynomials:

$$\nu H_n(\nu) - \frac{dH_n(\nu)}{d\nu} = H_{n+1}(\nu). \quad (\text{A.12})$$

The ratio of the non-Gaussian Edgeworth mass function to the Gaussian has the same structural form for either scaling:

$$\left. \frac{n_{\text{NG}}}{n_{\text{G}}} \right|_{\text{Edgeworth}} \approx 1 + \frac{F_1^{\text{h},f'}(M)}{F_0'(M)} + \frac{F_2^{\text{h},f'}(M)}{F_0'(M)} + \dots \quad (\text{A.13})$$

with the derivatives of each term  $F_s' = dF_s/dM$  for  $s \geq 1$ :

$$\begin{aligned} F_s^{\text{h},f'}(\nu) &= F_0' \sum_{\{k_m\}_h} \left\{ H_{s+2r} \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right. \\ &\quad \left. + H_{s+2r-1} \frac{\sigma}{\nu} \frac{d}{d\sigma} \left[ \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right] \right\} \\ F_s^{\text{f},f'}(\nu) &= F_0' \sum_{\{k_m\}_f} \left\{ H_{s+2} \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right. \\ &\quad \left. + H_{s+1} \frac{\sigma}{\nu} \frac{d}{d\sigma} \left[ \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right] \right\} \end{aligned} \quad (\text{A.14})$$

where the  $\{k_m\}$  again satisfy the relationships given below Eq.(A.11) and we have used

$$F_0' = \frac{e^{-\frac{\nu_c^2}{2}}}{\sqrt{2\pi}} \frac{d\sigma}{dM} \frac{\nu_c}{\sigma}. \quad (\text{A.15})$$

We have shared computer code to evaluate series for these two scalings at <http://www.slac.stanford.edu/~amantz/work/nongauss13/>.

## A.1 Other non-Gaussian mass functions

The mass function above has been shown to agree relatively well with  $N$ -body simulations of the local ansatz [45, 55], although there is some disagreement in the literature about just how well it agrees [67]. Some work has found evidence of a better fit if one shifts the collapse threshold [68], but our data are not sensitive to that level of detail. Ideally, one would like to use mass functions explicitly calibrated on simulations. To that end, Dalal et al. [20] proposed the following

$$\frac{dn}{dM} = \int dM_{\text{G}} \frac{dn}{dM_{\text{G}}} \frac{dP}{dM}(M_{\text{G}}) \quad (\text{A.16})$$

where  $dn/dM_G$  is the Gaussian mass function (which Dalal et al. took to be Jenkins et al. [48]) and  $dP/dM(M_G)$  is the probability distribution describing how a Gaussian halo of mass  $M_G$  maps into a non-Gaussian halo with mass  $M$ . Dalal et al. gave fitting formula for this as a Gaussian distribution with:

$$\begin{aligned} \left\langle \frac{M}{M_G} \right\rangle - 1 &= 1.3 \times 10^{-4} f_{\text{NL}} \sigma_8 \sigma(M_G, z)^{-2} \\ \text{var} \left( \frac{M}{M_G} \right) &= 1.4 \times 10^{-4} (f_{\text{NL}} \sigma_8)^{0.8} \sigma(M_G, z)^{-1}. \end{aligned} \quad (\text{A.17})$$

For a fixed value of  $f_{\text{NL}}$ , this mass function deviates less from the Gaussian than the LMSV mass function does. This is an interesting approach, but has not been tried for non-local non-Gaussianity.

For reference, we collect here the various non-Gaussian mass functions that exist in the literature. The mass function can be parametrized as

$$\frac{dn}{dM}(M, z) = f(\sigma) \frac{\bar{\rho}}{M} \frac{d \ln[\sigma^{-1}(M, z)]}{dM}, \quad (\text{A.18})$$

where  $f(\sigma)$  will differ between various models. For reference, we compile here  $f(\sigma)$  for the (Gaussian) Tinker mass function [33] ( $f_{\text{T}, \Delta=300}$ ), the LMSV mass function [30] truncated at first order ( $f_{\text{LMSV}, 1}$ ), the Dalal et al. [20] mass function ( $f_{\text{Dalal}}$ ), and the LMSV mass function keeping a large number ( $n$ ) of terms with either hierarchical ( $f_{\text{LMSV}, n}^{\text{h}}$ ) or feeder ( $f_{\text{LMSV}, n}^{\text{f}}$ ) scaling for the cumulants:

$$\begin{aligned} f_{\text{T}, \Delta=300}(\sigma) &= A \left[ \left( \frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-c/\sigma^2}, \\ f_{\text{LMSV}, 1} &= f_{\text{T}, \Delta=300} \left[ 1 + \frac{F_1^{\text{h}, \text{f}'}(M)}{F_0'(M)} \right], \\ &= f_{\text{T}, \Delta=300} \left[ 1 + H_3 \frac{\mathcal{M}_3}{3!} + H_2 \frac{\sigma}{\nu} \frac{d\mathcal{M}_3}{d\sigma} \right], \\ f_{\text{Dalal}}(\sigma_M) &= \left[ \frac{d \ln[\sigma^{-1}(M, z)]}{dM} \right]^{-1} \int dM_G f_{\text{T}, \Delta=300}(\sigma_{M_G}) \frac{d \ln[\sigma^{-1}(M_G, z)]}{dM_G} \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma_{M_G}} \exp \left[ -\frac{(\frac{M}{M_G} - \langle \frac{M}{M_G} \rangle)^2}{2\sigma_{M/M_G}^2} \right], \\ f_{\text{LMSV}, n}^{\text{h}, \text{f}} &= f_{\text{T}, \Delta=300} \left[ 1 + \sum_{s=1}^n \frac{F_s^{\text{h}, \text{f}'}}{F_0'} \right]. \end{aligned} \quad (\text{A.19})$$

In general, anywhere we have written  $f_{\text{T}, \Delta=300}$  in the non-Gaussian mass functions, one can substitute any other Gaussian model. For models with hierarchical scaling, it is common to use the skewness parameter  $S_3 = \mathcal{M}_3/\sigma$ . With that definition, it is clear that the second line is equivalent to the expression used for the LMSV mass function in [7, 8],

$$f_{\text{LMSV}, 1} = f_{\text{T}, \Delta=300} \left[ 1 + \frac{1}{6} \frac{\sigma^2}{\delta_c} \left[ S_3 \left( \frac{\delta_c^4}{\sigma^4} - 2 \frac{\delta_c^2}{\sigma^2} - 1 \right) + \frac{1}{6} \frac{dS_3}{d \ln \sigma} \left( \frac{\delta_c^2}{\sigma^2} - 1 \right) \right] \right]. \quad (\text{A.20})$$

## References

- [1] J. P. Henry, A. E. Evrard, H. Hoekstra, A. Babul, and A. Mahdavi, *The X-ray Cluster Normalization of the Matter Power Spectrum*, *Astrophys.J.* **691** (2009) 1307–1321, [[arXiv:0809.3832](#)].
- [2] A. Vikhlinin, A. Kravtsov, R. Burenin, H. Ebeling, W. Forman, *et. al.*, *Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints*, *Astrophys.J.* **692** (2009) 1060–1074, [[arXiv:0812.2720](#)].
- [3] A. Mantz, S. W. Allen, D. Rapetti, and H. Ebeling, *The Observed Growth of Massive Galaxy Clusters I: Statistical Methods and Cosmological Constraints*, *Mon.Not.Roy.Astron.Soc.* **406** (2010) 1759–1772, [[arXiv:0909.3098](#)].
- [4] J. Truemper, *ROSAT - A new look at the X-ray sky*, *Science.* **260** (1993) 1769–1771.
- [5] A. Mantz, S. W. Allen, H. Ebeling, D. Rapetti, and A. Drlica-Wagner, *The Observed Growth of Massive Galaxy Clusters II: X-ray Scaling Relations*, *Mon.Not.Roy.Astron.Soc.* **406** (2010) 1773–1795, [[arXiv:0909.3099](#)].
- [6] C. Cunha, D. Huterer, and O. Dore, *Primordial non-Gaussianity from the covariance of galaxy cluster counts*, *Phys.Rev.* **D82** (2010) 023004, [[arXiv:1003.2416](#)].
- [7] A. Pillepich, C. Porciani, and T. H. Reiprich, *The X-ray cluster survey with eROSITA: forecasts for cosmology, cluster physics, and primordial non-Gaussianity*, *Mon.Not.Roy.Astron.Soc.* **422** (2012) 44–69, [[arXiv:1111.6587](#)].
- [8] D. S. Mak and E. Pierpaoli, *Constraints on Non-Gaussianity from Sunyaev–Zeldovich Cluster Surveys*, *Phys.Rev.* **D86** (2012) 123520, [[arXiv:1210.7276](#)].
- [9] C. Bennett, D. Larson, J. Weiland, N. Jarosik, G. Hinshaw, *et. al.*, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, [[arXiv:1212.5225](#)].
- [10] **Planck** Collaboration, P. Ade *et. al.*, *Planck 2013 Results. XXIV. Constraints on primordial non-Gaussianity*, [[arXiv:1303.5084](#)].
- [11] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, *Constraints on local primordial non-Gaussianity from large scale structure*, *JCAP* **0808** (2008) 031, [[arXiv:0805.3580](#)].
- [12] J.-Q. Xia, A. Bonaldi, C. Baccigalupi, G. De Zotti, S. Matarrese, *et. al.*, *Constraining Primordial Non-Gaussianity with High-Redshift Probes*, *JCAP* **1008** (2010) 013, [[arXiv:1007.1969](#)].
- [13] J.-Q. Xia, C. Baccigalupi, S. Matarrese, L. Verde, and M. Viel, *Constraints on Primordial Non-Gaussianity from Large Scale Structure Probes*, *JCAP* **1108** (2011) 033, [[arXiv:1104.5015](#)].
- [14] T. Giannantonio, A. J. Ross, W. J. Percival, R. Crittenden, D. Bacher, *et. al.*, *Improved Primordial Non-Gaussianity Constraints from Measurements of Galaxy Clustering and the Integrated Sachs-Wolfe Effect*, [[arXiv:1303.1349](#)].
- [15] B. Benson, T. de Haan, J. Dudley, C. Reichardt, K. Aird, *et. al.*, *Cosmological Constraints from Sunyaev-Zel’dovich-Selected Clusters with X-ray Observations in the First 178 Square Degrees of the South Pole Telescope Survey*, *Apj*, **763**, **147** (2013) [[arXiv:1112.5435](#)].
- [16] R. Williamson, B. Benson, F. High, K. Vanderlinde, P. Ade, *et. al.*, *An SZ-selected sample of the most massive galaxy clusters in the 2500-square-degree South Pole Telescope survey*, *Astrophys.J.* **738** (2011) 139, [[arXiv:1101.1290](#)].
- [17] A. Mana, T. Giannantonio, J. Weller, B. Hoyle, G. Huetsi, *et. al.*, *Combining clustering and abundances of galaxy clusters to test cosmology and primordial non-Gaussianity*, [[arXiv:1303.0287](#)].

- [18] D. S. Salopek and J. R. Bond, *Nonlinear evolution of long wavelength metric fluctuations in inflationary models*, *Phys. Rev.* **D42** (1990) 3936–3962.
- [19] E. Komatsu and D. N. Spergel, *Acoustic signatures in the primary microwave background bispectrum*, *Phys.Rev.* **D63** (2001) 063002, [[astro-ph/0005036](#)].
- [20] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, *The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects*, *Phys.Rev.* **D77** (2008) 123514, [[arXiv:0710.4560](#)].
- [21] C. Fedeli, C. Carbone, L. Moscardini, and A. Cimatti, *The clustering of galaxies and galaxy clusters: constraints on primordial non-Gaussianity from future wide-field surveys*, *Mon.Not.Roy.Astron.Soc.* **414** (2011) 1545–1559, [[arXiv:1012.2305](#)].
- [22] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich, *Constraining primordial non-Gaussianity with future galaxy surveys*, *Mon.Not.Roy.Astron.Soc.* **422** (2012) 2854–2877, [[arXiv:1109.0958](#)].
- [23] A. R. Pullen and C. M. Hirata, *Systematic effects in large-scale angular power spectra of photometric quasars and implications for constraining primordial nongaussianity*, [arXiv:1212.4500](#).
- [24] D. Huterer, C. E. Cunha, and W. Fang, *Calibration errors unleashed: effects on cosmological parameters and requirements for large-scale structure surveys*, [arXiv:1211.1015](#).
- [25] P. Creminelli and M. Zaldarriaga, *Single field consistency relation for the 3-point function*, *JCAP* **0410** (2004) 006, [[astro-ph/0407059](#)].
- [26] C. Wagner and L. Verde, *N-body simulations with generic non-Gaussian initial conditions II: Halo bias*, *JCAP* **1203** (2012) 002, [[arXiv:1102.3229](#)].
- [27] N. Agarwal, S. Ho, and S. Shandera, *In preparation*, .
- [28] N. Barnaby, R. Namba, and M. Peloso, *Observable non-gaussianity from gauge field production in slow roll inflation, and a challenging connection with magnetogenesis*, *Phys.Rev.* **D85** (2012) 123523, [[arXiv:1202.1469](#)].
- [29] B. P. Koester, T. A. McKay, J. Annis, R. H. Wechsler, A. Evrard, L. Bleem, M. Becker, D. Johnston, E. Sheldon, R. Nichol, C. Miller, R. Scranton, N. Bahcall, J. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. Kleinman, J. Krzesinski, D. Long, A. Nitta, D. P. Schneider, S. Sneddin, W. Voges, and D. York, *A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey*, *ApJ* **660** (May, 2007) 239–255, [[astro-ph/](#)].
- [30] M. LoVerde, A. Miller, S. Shandera, and L. Verde, *Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures*, *JCAP* **0804** (2008) 014, [[arXiv:0711.4126](#)].
- [31] N. Barnaby and S. Shandera, *Feeding your Inflaton: Non-Gaussian Signatures of Interaction Structure*, *JCAP* **1201** (2012) 034, [[arXiv:1109.2985](#)].
- [32] S. Shandera, A. L. Erickcek, P. Scott, and J. Y. Galarza, *Number Counts and Non-Gaussianity*, [arXiv:1211.7361](#).
- [33] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, *et. al.*, *Toward a halo mass function for precision cosmology: The Limits of universality*, *Astrophys.J.* **688** (2008) 709–728, [[arXiv:0803.2706](#)].
- [34] **WMAP** Collaboration, E. Komatsu *et. al.*, *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys.J.Suppl.* **192** (2011) 18, [[arXiv:1001.4538](#)].
- [35] N. Jarosik, C. Bennett, J. Dunkley, B. Gold, M. Greason, *et. al.*, *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results*, *Astrophys.J.Suppl.* **192** (2011) 14, [[arXiv:1001.4744](#)].

- [36] X. Chen and Y. Wang, *Quasi-Single Field Inflation and Non-Gaussianities*, *JCAP* **1004** (2010) 027, [[arXiv:0911.3380](#)].
- [37] H. Ebeling, A. Edge, H. Bohringer, S. Allen, C. Crawford, *et. al.*, *The ROSAT Brightest Cluster Sample (BCS) - I. The compilation of the sample and the cluster log N-log S distribution*, *Mon.Not.Roy.Astron.Soc.* **301** (1998) 881–914, [[astro-ph/9812394](#)].
- [38] H. Bohringer, P. Schuecker, L. Guzzo, C. Collins, W. Voges, *et. al.*, *The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy Cluster Survey. 5. The Cluster catalogue*, *Astron.Astrophys.* **425** (2004) 367–383, [[astro-ph/0405546](#)].
- [39] H. Ebeling, A. Edge, A. Mantz, E. Barrett, J. Henry, *et. al.*, *The X-ray brightest clusters of galaxies from the Massive Cluster Survey*, *Mon.Not.Roy.Astron.Soc.* **407** (2010) 83–93, [[arXiv:1004.4683](#)].
- [40] S. Allen, D. Rapetti, R. Schmidt, H. Ebeling, G. Morris, *et. al.*, *Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters*, *Mon.Not.Roy.Astron.Soc.* **383** (2008) 879–896, [[arXiv:0706.0033](#)].
- [41] A. G. Riess, L. Macri, S. Casertano, M. Sosey, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, R. Chornock, and D. Sarkar, *A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder*, *ApJ* **699** (July, 2009) 539–563, [[arXiv:0905.0695](#)].
- [42] D. Kirkman, D. Tytler, N. Suzuki, J. M. O’Meara, and D. Lubin, *The Cosmological baryon density from the deuterium to hydrogen ratio towards QSO absorption systems: D/H towards Q1243+3047*, *Astrophys.J.Suppl.* **149** (2003) 1, [[astro-ph/0302006](#)].
- [43] B. Sartoris, S. Borgani, C. Fedeli, S. Matarrese, L. Moscardini, *et. al.*, *The potential of X-ray cluster surveys to constrain primordial non-Gaussianity*, *Mon.Not.Roy.Astron.Soc.* **407** (2010) 2339–2354, [[arXiv:1003.0841](#)].
- [44] M. Oguri, *Self-Calibrated Cluster Counts as a Probe of Primordial Non-Gaussianity*, *Phys.Rev.Lett.* **102** (2009) 211301, [[arXiv:0905.0920](#)].
- [45] A. Pillepich, C. Porciani, and O. Hahn, *Halo mass function and scale-dependent bias from N-body simulations with non-Gaussian initial conditions*, *Mon.Not.Roy.Astron.Soc.* **402** (2008) 191–206, [[arXiv:0811.4176](#)].
- [46] R. K. Sheth, H. J. Mo, and G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, *Mon. Not. Roy. Astron. Soc.* **323** (2001) 1, [[astro-ph/9907024](#)].
- [47] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro, *Precision determination of the mass function of dark matter halos*, *Astrophys. J.* **646** (2006) 881–885, [[astro-ph/0506395](#)].
- [48] A. Jenkins, C. Frenk, S. D. White, J. Colberg, S. Cole, *et. al.*, *The Mass function of dark matter halos*, *Mon.Not.Roy.Astron.Soc.* **321** (2001) 372, [[astro-ph/0005260](#)].
- [49] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber, *The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, *ApJ* **724** (Dec., 2010) 878–886, [[arXiv:1001.3162](#)].
- [50] S. Majumdar and J. J. Mohr, *Self calibration in cluster studies of dark energy: Combining the cluster redshift distribution, the power spectrum and mass measurements*, *Astrophys.J.* **613** (2004) 41–50, [[astro-ph/0305341](#)].
- [51] H.-Y. Wu, E. Rozo, and R. H. Wechsler, *Annealing a Follow-up Program: Improvement of the Dark Energy Figure of Merit for Optical Galaxy Cluster Surveys*, *Astrophys.J.* **713** (2010) 1207–1218, [[arXiv:0907.2690](#)].
- [52] E. S. Sheldon, D. E. Johnston, R. Scranton, B. P. Koester, T. A. McKay, H. Oyaizu,

- C. Cunha, M. Lima, H. Lin, J. A. Frieman, R. H. Wechsler, J. Annis, R. Mandelbaum, N. A. Bahcall, and M. Fukugita, *Cross-correlation Weak Lensing of SDSS Galaxy Clusters. I. Measurements*, *ApJ* **703** (Oct., 2009) 2217–2231, [[arXiv:0709.1153](#)].
- [53] D. E. Johnston, E. S. Sheldon, R. H. Wechsler, E. Rozo, B. P. Koester, J. A. Frieman, T. A. McKay, A. E. Evrard, M. R. Becker, and J. Annis, *Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation*, *ArXiv e-prints* (Sept., 2007) [[arXiv:0709.1159](#)].
- [54] J. Fergusson, D. Regan, and E. Shellard, *Optimal Trispectrum Estimators and WMAP Constraints*, [arXiv:1012.6039](#).
- [55] M. LoVerde and K. M. Smith, *The Non-Gaussian Halo Mass Function with fNL, gNL and tauNL*, *JCAP* **1108** (2011) 003, [[arXiv:1102.1439](#)].
- [56] S. Adhikari, L. Book, N. Dalal, and S. Shandera, *In preparation*, .
- [57] A. von der Linden, M. T. Allen, D. E. Applegate, P. L. Kelly, S. W. Allen, *et. al.*, *Weighing the Giants I: Weak Lensing Masses for 51 Massive Galaxy Clusters - Project Overview, Data Analysis Methods, and Cluster Images*, [arXiv:1208.0597](#).
- [58] D. E. Applegate, A. von der Linden, P. L. Kelly, M. T. Allen, S. W. Allen, *et. al.*, *Weighing the Giants III: Methods and Measurements of Accurate Galaxy Cluster Weak-Lensing Masses*, [arXiv:1208.0605](#).
- [59] **Dark Energy Survey** Collaboration, T. Abbott *et. al.*, *The dark energy survey*, [astro-ph/0510346](#).
- [60] A. Merloni, P. Predehl, W. Becker, H. Bohringer, T. Boller, *et. al.*, *eROSITA Science Book: Mapping the Structure of the Energetic Universe*, [arXiv:1209.3114](#).
- [61] W. H. Press and P. Schechter, *Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation*, *Astrophys. J.* **187** (1974) 425–438.
- [62] W. A. Chiu, J. P. Ostriker, and M. A. Strauss, *Using cluster abundances and peculiar velocities to test the gaussianity of the cosmological density field*, *Astrophys. J.* **494** (1998) 479–490, [[astro-ph/9708250](#)].
- [63] J. Robinson and J. E. Baker, *Evolution of the cluster abundance in non-gaussian models*, *Mon. Not. Roy. Astron. Soc.* **311** (2000) 781, [[astro-ph/9905098](#)].
- [64] S. Matarrese, L. Verde, and R. Jimenez, *The Abundance of high-redshift objects as a probe of non-Gaussian initial conditions*, *Astrophys.J.* **541** (2000) 10, [[astro-ph/0001366](#)].
- [65] V. V. Petrov, *Sums of independent random variables*. Springer-Verlag, Berlin, New York, 1975.
- [66] V. V. Petrov, *Limit theorems of probability theory: sequences of independent random variables*, vol. 4 of *Oxford studies in probability*. Clarendon Press, Oxford / Oxford University Press, New York, 1995.
- [67] X. Kang, P. Norberg, and J. Silk, *Can large-scale structure probe CMB-constrained non-Gaussianity?*, *Mon.Not.Roy.Astron.Soc.* **376** (2007) 343–347, [[astro-ph/0701131](#)].
- [68] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, *et. al.*, *Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations*, *Mon.Not.Roy.Astron.Soc.* **398** (2009) 321–332, [[arXiv:0902.2013](#)].